Filed on behalf of: SENIOR PARTY: Advanced Life Science Institute, Inc.

By: Eugene Lieberstein, Esq.
David Einhorn, Esq.
Joe Manak, Esq.
ANDERSON KILL & OLICK
1251 Avenue of the Americas
New York, NY 10020
Tel: 212-278-1359
Fax: 212-278-1733

Filed: February 2, 2007

BOX INTERFERENCE

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES

Advanced Life Science Institute, Inc.,
Senior Party
(Patent No. 5,976,832;
Inventors: Jiro Hitomi; Ken Yamaguchi; Tokujiro Yamamura; and
Tatsuki Kimura).

v.

Human Genome Sciences, Inc.,
Junior Party
(Application No. 08/761,289,
Inventors: Jian Ni; Guo-Liang Yu; Pedro Alfonso; Reiner Gentz; and
Jeffrey Y. Su),

Patent Interference 105,501 (RES)
(Technology Center 1600)

ALS MOTION NO. 2

Second Weber Declaration
(Exhibit No. 1010)

ALS EXHIBIT 1010
ALS, Inc. v. HGS, Inc.
Interference Case 105,501

SECOND DECLARATION OF DAVID J. WEBER, PHD

I, David J. Weber, a citizen of the United States of America, hereby declare and state:

1. I completed Post Doctoral Training in the Department of Biological Chemistry at Johns Hopkins School of Medicine from 1988 until 1993. I have a degree in Biological Chemistry (Ph.D.), which was conferred upon me by The University of North Carolina – Chapel Hill, Chapel Hill North Carolina in 1988. I also have a degree in Chemistry (B.S.), which was conferred upon me by Muhlenberg College in Allentown, Pennsylvania, in 1984.

2. I have been employed by the University of Maryland School of Medicine, Baltimore, Maryland since 1993. I was hired as a tenure-track Assistant Professor and was promoted to Associate Professor in 1999 and awarded tenure. In 2004, I was promoted to Full Professor (until present), and in 2006 I was Acting Chair of The Department of Biochemistry and Molecular Biology.

3. I started working on S100 proteins, the subject of the present Interference, in 1993, as an independent Principle Investigator and have been funded continuously by multiple funding sources since that time to do research studying the structure and function of S100 proteins. Specifically, the funding sources for the research in my laboratory included grants from The University of Maryland School of Medicine (i.e. start-up package) and numerous peer-reviewed grants from the National Institutes of Health (NIH), The National Science Foundation (NSF), and the American Cancer Society (ACS). Currently, I am the Principle Investigator of two NIH grants, one ACS grant, and co-Principle Investigator on another NIH grant. Students and post

doctoral fellows in my laboratory are also funded via training grants (NIH, ACS), and I am their mentor. I am an author on a total of one-hundred and forty seven publications including sixty-two peer-reviewed articles, ten invited book chapters/review papers, and seventy-five published abstracts most of which are written on topics about the structure and function of various S100 proteins.

4. I am a referee for eighteen scientific journals in the field of Biochemistry & Molecular Biology and I have been on scientific panels to review grant applications for the National Institutes of Health, The National Science Foundation, The American Cancer Society, The American Chemical Society, and The Israel Science Foundation and for Searle Scholar selections.

5. A copy of my curriculum vita is attached as an addendum separately identified as Exhibit 1010.

6. I have been asked to give my opinion as to whether the two calcium binding proteins encoded by DNA, which correspond respectively to the two amino acid sequences SEQ ID NO. 19 and SEQ ID NO. 20 as defined in claim 1 of US Patent No. 5,976,832 (the '832 patent), designated as corresponding to the Count in this Interference, represented novel proteins to one skilled in the field of biological chemistry in 1995 in relation to all other known members in the S100 family of calcium binding proteins at the time of the filing of the '832 patent and whether these two proteins are independent chemical entities distinct from one another to the extent that it would not have been obvious to one of ordinary skill in the art at the time of the filing of the '832 patent to derive one of these two proteins from the other.

7. As has been explained to me by counsel, if each of the above calcium binding proteins were novel to one skilled in this art, at the time of the filing of

the '832 patent, and if each is independent and distinct with respect to one another in that neither of the proteins would have been obvious from knowledge of the other at the time of the filing of the '832 patent, then they constitute patentably distinct chemical entities and accordingly claim 1 of the '832 patent defines more than one patentably distinct invention.

8. In my opinion and for the reasons I set out below the two proteins encoded by DNA corresponding to the two amino acid sequences SEQ ID NO. 19 and SEQ ID NO. 20, as defined in claim 1 of the '832 patent, were novel at the time of the filing of the '832 patent and are both independent, distinct and non-obvious with respect to one another.

Introduction to S100 proteins.

9. This interference concerns certain proteins belonging to what is known as the S100 group or "family" of proteins. S100 proteins constitute a family of Ca^{2+}-binding proteins with twenty-four unique full-length amino acid sequences discovered to date. At the time of the filing of the '832 patent in 1995, approximately half as many S100 proteins were known.

10. The sequence listing Seq ID no. 20 in claim 1 of the '832 patent is an amino acid sequence representing a single protein with the updated name: human S100A12. The sequence listing Seq ID no. 19 in claim 1 of the '832 patent is an amino acid sequence representing a single protein derived from bovine given the name CAAF1 by Hitomi et al in the '832 patent and in the following publication of Hitomi: Hitomi et al. "A novel calcium-binding protein in amniotic fluid, CAAF1: its molecular cloning and tissue distribution," Journal of Cell Science, Vol. 109, pp. 805-815 (1996)

(hereinafter "Hitomi 1996"). Exhibit 1004. This publication "Hitomi 1996" correctly recognizes protein CAAF1 from bovine (Seq ID no. 19) as being a new and novel S100 protein.

11. It should be understood that although sequence listing Seq ID no. 20 was discovered based upon sequence listing Seq ID no. 19 in the '832 patent; the sequence listing Seq ID no. 20 is not the human counterpart of sequence listing Seq ID no. 19. This is self evident from the substantial differences in the protein sequence for Seq ID no. 19 as compared to Seq ID no. 20 as shown below. If Seq ID no. 20 was the human counterpart to Seq ID no. 19, we would see only a very few amino acid differences between the two proteins (<5). The fact that we see a very large number of differences (total of 31), demonstrates that Seq ID no. 19 and Seq ID no. 20 are completely distinct and independent proteins.

12. Amino acid sequence Seq. ID no.: 19 is as follows:

TKLE**D**HLEGI**I**NIFHQYSVR**V**GHFDTL**N**K**R**ELKQL**I**TKEL**P**K**T**L**Q**N**T**
KD**Q**P**T**IDK**I**FQ**D**LDA**D**K**D**G**A**V**S**F**E**EF**V**V**L**V**S**R**V**LK**T**AH**I**D**I**HKE

13. Amino acid sequence Seq. ID No.: 20 is as follows:

TKLEEHLEGIVNIFHQYSVRKGHFDTLSKGELKQLLTKELANTIKNIKDKA
VIDEIFQGLDANQDEQVDFQEFISLVAIALKAAHYHTHKE

(For purposes of comparison the residues in Seq ID No. 19 which are different from those in Seq ID No. 20 are shown bold and underlined.)

14. In a general sense all S100 proteins, including S100A12 (Seq ID no. 20 in claim 1 of the '832 patent) and CAAF1 (Seq ID no. 19 in claim 1 of the '832 patent), have regions of amino acid residues that are similar to all other S100 proteins. These regions of identity and similarity pertain to the ability of S100 proteins to bind calcium. However, these two proteins differ in other important regions of each protein

sequence. It is these regions of diversity, which are exposed upon binding calcium, that enables an S100 protein to bind another specific protein, termed a target protein. In fact, most S100 proteins have multiple protein targets, including S100A12, which has several. It is the interaction with these specific target proteins, which is important for a specific biological effect of a specific S100 protein. Thus, there are similarities and differences in the biological functions of specific S100 proteins as well as differences in their amino acid and polynucleotide sequences.

15. While comparing gene sequences is fairly straightforward, it is necessary to examine their corresponding proteins to understand how the expressed gene product biologically functions. The 3D structure of the S100 protein, which is encoded by DNA, and its complexes with target proteins is critically important to understand how the protein functions biologically.

Novelty

16. At the time of filing of the '832 patent neither SEQ ID No. 19 or SEQ ID No. 20 was known to the public inclusive of those skilled in the art.

17. Since there are 31 amino acid residues in Seq ID No. 19, which are different than those in Seq ID No. 20, the two sequences are clearly independent and novel relative to one another. In addition, and as explained in greater detail hereafter, these proteins have distinctly different biological properties, which clearly makes them non-obvious over one another.

Obviousness

18. With respect to the issue of obviousness, I have determined that one of ordinary skill in the art at the time of the filing of the '832 patent would not have had a reasonable expectation of success at arriving at an "isolated and purified DNA

encoding a calcium binding protein that possess calcium binding activity comprising an amino acid sequence, which is identical to the amino acid sequence listed in SEQ ID NO: 19 . . ." of the '832 patent with the knowledge of the SEQ ID NO. 20 or vice versa.

The Ordinary Skilled Worker

19. In 1995, an entry level technician with a BS or BA or RN degree would typically have less than six years of experience on average. A person with a PhD and/or a MS in specified fields such as Biochemistry, Molecular Biology, Genetics, Immunology, Microbiology, Neuroscience, Chemistry, and other recognized fields in Biological and Life Sciences would also be representative of an ordinary skilled worker. The MS scientist would have six years of experience on average, and the PhD would have eight to ten years of experience on average. A person with a medical degree, MD, could also be considered an ordinary skilled worker in 1995 and on average would have five to seven year of experience. A person with an MD/PhD dual degree would also be considered an ordinary worker, and like the PhD, would have eight to ten years of experience on average.

20. Specific procedures involving oligonucleotide probes, gene cloning, expression vectors, host cells, and planning for protein purification/production, are outlined in a very specific manner in a laboratory manual from the Cold Spring Harbor entitled: "Molecular Cloning; A Laboratory Manual (second edition)" co-authored by Sambrook, Fritsch, and Maniatis (Editors, N. Irwin, N. Ford, C. Nolan, M. Ferguson, M. Ockler). Cold Spring Harbor Laboratory Press (1989). Exhibit 1005. Such methodologies were performed by a large number of ordinary skilled workers in 1995.

21. A quote from Sambrook *et al.* states "The synthesis and cloning of cDNA are still not easy, but as a consequence of a wide range and technical and

theoretical advances, cDNA cloning is now well within the range of any competent laboratory". Also as pointed out in this same manual, "the synthesis and cloning of cDNA requires a careful analysis of available methods, vectors, and screening procedures that offer the best chance of success; which combination of such factors that is most appropriate for a particular cloning problem depends on the information and reagents that are available." Exhibit 1005.

22. In 1995, an ordinary skilled worker would not have a reasonable expectation of success to arrive at a specific clone (DNA and protein sequence) with an exact sequence (i.e., Seq ID no. 19 or Seq ID no. 20) if they had knowledge of the protein sequence of the other (i.e., Seq. ID. No. 20 or Seq. ID no. 19) since each is a protein of high homology but does not have the exact identity of the other.

23. As demonstrated below, knowing the amino acid sequence of one S100 protein (e.g., Seq ID no. 20 or Seq ID no. 19), using techniques available in 1995, was not sufficient for the ordinary skilled worker to have reasonable expectation of attaining another S100 protein with a different, but defined sequence.

24. As demonstrated above the Seq. ID no. 20 and Seq. ID no. 19 are different chemical entities. As I will demonstrate below, from the techniques and material available in 1995, more than one DNA sequence would be obtained when only the amino acid sequence of one S100 protein were known and furthermore, the exact DNA and amino acid sequences discovered with such methods were not predictable with any certainty.

25. Because of the lack of predictability of sequences found using such experimentation, it is clear that a skilled worker would not have a reasonable expectation of attaining one S100 protein with a different, but defined sequence using

techniques available in 1995 even if he/she had the amino acid sequence of another S100 protein such as the Seq. ID no. 20 or the Seq. ID no. 19 in hand.

26. In general, proteins of the S100 family are highly homologous in their two calcium-binding EF-hand domains and in regions of the protein that are necessary to form the X-type four helix bundle at the dimer interface.

27. Another similarity is that nearly all S100 proteins exhibit a conformational change that exposes a hydrophobic patch via which S100 proteins interact with other proteins referred to as target proteins.

28. However, S100 family members lack sequence homology in the 'hinge region' (loop 2; residues 40-50 or thereabouts depending on the specific S100 protein) and in the C-terminal loop (at or nearby residue 78, depending on the S100 protein, to the C-terminus). Such differences in sequence in these two regions, and in other regions in some cases, explains why specificity in target protein binding is observed among many S100 family members.

29. For these reasons, it is not surprising that the two S100 protein sequences i.e., Seq ID no. 19 and Seq ID no. 20 have regions of high homology (i.e. EF-hand calcium-binding domains and dimerization domain), but also contain regions of diversity (loop 2 termed the hinge region; the C-terminal loop). Such differences are typically found when S100 proteins are compared.

30. Below are the specific sequences with the residues showing different amino acids illustrated with an underline/bold of the 1 letter amino acid code:

#19
TKLE**D**HLEGI**I**NIFHQYSVR**V**GHFDTL**N**K**R**ELKQL**I**TKEL**P**K**T**L**Q**N**T**
KD**Q**P**T**ID**K**IFQ**D**LDA**D**K**D**G**A**V**S**F**E**EF**V**V**L**V**S**R**V**LK**T**AH**I**D**I**HKE

#20
TKLEEHLEGIVNIFHQYSVRKGHFDTLSKGELKQLLTKELANTIKNI
KDKAVIDEIFQGLDANQDEQVDFQEFISLVAIALKAAHYHTHKE

31. As explained earlier , 31 out of 91 residues (>30%; underlined and in bold above) are not identical between the amino acid residues.

32. Because multiple 3 letter codon usages can code for a single amino acid, the nucleotide sequence encoding these two proteins (Seq ID nos. 20 and 19) are even that much more diverse when compared.

33. Another important observation is that the majority of the differences observed in the two protein sequences (Seq ID Nos. 20 and 19) are in the hinge region (loop 2; residues 40-53) in the middle of the protein and at the C-terminus of the protein (residues 75 to the C-terminus). It is these regions of S100 proteins that typically have the least sequence homology and, it is these regions that interact with specific target proteins in S100-target protein interactions to generate specific biological effects. It can also be inferred that these two proteins interact differently with target proteins and indeed they likely interact with different target proteins all together. It may also be inferred, that different target protein interactions between these two S100 proteins (Seq ID no. 20 and Seq ID no. 19) would correspondingly mean that the two proteins will likely have different functions in a cell.

The techniques available in 1995 to clone, sequence, and engineer a gene and its corresponding recombinant protein demonstrate that the ordinary skilled worker in 1995 would not have had a reasonable chance of success in arriving at either Seq. ID No. 19 or Seq. ID No. 20 with the knowledge of the other.

34. In 1995, the cloning of a gene started with knowledge about the protein sequence of the gene product.

35. When a biological activity of interest was identified in 1995, researchers would do purification steps starting with tissues and/or specific cells and monitor the activity of interest. As they removed protein components that did not contribute to this activity, the ratio of activity/total protein would increase. This ratio of activity/total protein is termed "specific activity". As an example, for calcium-binding proteins such as the S100 family members, an activity that could be measured was calcium-binding using $^{45}Ca^{2+}$ overlays. Once they had a purified protein, as judged from multiple criteria, the specific activity would no longer change upon doing additional purification steps. At this point, the protein would be considered pure and the exact protein, but not DNA sequence, could be determined. The conclusion could then be made that this purified protein with a defined amino acid sequence contributes to the activity measured in the tissue or cell. A variety of chromatography experiments would have to be tested on a trial and error basis in a systematic way including testing buffering conditions, pH, salt concentrations etc. Once a step was shown to remove a large amount of contaminants, then another chromatography test would be performed, and so on, until the specific activity no longer changed and the protein was shown to be pure using a variety of methods (SDS-PAGE, mass spec, etc). Such experiments have been done with several S100 proteins including S100A12 (Seq ID No. 20) and CAAF1 (Seq ID No. 19).

36. To clone a gene in 1995, with the intent of obtaining the active protein, the first step would be to create a cDNA library from cells or tissue that had mRNA with the gene of interest. It was sometimes necessary to enrich the mRNA prior to cDNA production, and it was always important to check the integrity of the mRNA in some manner (i.e. translation assays, size, ability to make gene of interest etc) prior to

cDNA production, which used pure reverse transcriptase (i.e. free of contaminating RNAase, DNAase enzymes).

37. An oligo(dT) primer was usually used to prime DNA synthesis by binding to Eukaryotic poly(A) tracts at the 3' of the cellular mRNA molecules in Eukaryotic cells. Sometimes it was beneficial to pre-treat the mRNA to remove RNA secondary structure and vary reaction temperatures; it was also beneficial to have RNAse inhibitors included. Typically, the second cDNA strand was made using DNA polymerase I after denaturing the DNA/RNA hybrid. If RNA:DNA hybrid denaturation was not sufficient, the second DNA strand could also be prepared by having the first strand treated with a terminal transferase to add a poly(G) tail (for example) for priming. PCR methods were then becoming available at this time to generate cDNA libraries if mRNAs were no larger than 1-2 kb in length. All such methodologies herein required optimizing solution conditions, reaction temperatures, reaction times etc in a trial an error manner.

38. While the foregoing procedures were known in the art in 1995, they were not simply a routine matter to carry out because each of the steps could have multiple strategies, and since each particular step included methods that were not always full-proof due to the nature of the materials involved. Thus, systematic trial and error experimentation with knowledge of multiple strategies for arriving at the next step were required to generate a good cDNA library. Particular attention to the quality of the mRNA at the early stage would have a large impact on the quality of the cDNA library, and ultimately one's ability to clone a gene from such a library.

39. Next, bacteriophage vectors (i.e. gamma-gt10, gamma-gt11, gamma-gt18 to gamm-gt23, gamma-ORF8, etc) were used in 1995 to amplify and store

the cDNA libraries. Libraries such as these could be stored indefinitely and probed using nucleic acid and antibody probes. There were several methods to select a particular vector with each having its own advantages and disadvantages, and as described before required trial and error and careful consideration with the choice of procedures depending on the materials chosen and/or the physical characteristics of the gene itself.

40. In 1995, there were two major ways to screen cDNA libraries for a clone of interest including: (1) nucleic acid hybridization and (2) immunological detection of specific antigens. Nucleic acid hybridization was the most commonly used technique and the most reliable method for screening clones of interest.

41. Once a cDNA library was available, which was derived from a tissue containing the gene product of interest, the use of polymerase chain reaction ("PCR") type cloning was most widely used to amplify this DNA.

42. The drawback of the PCR method was that it was not always full-proof due to the intrinsic error rates of the enzyme reagents used (i.e. the Taq polymerase used to effectuate PCR, incorporates the wrong nucleotide at a rate of 1:10,000; giving and overall error of 0.25% with 30 cycles). For this reason, it was imperative to fully sequence any DNA obtained using PCR to make sure there were no mutations produced via PCR.

43. The choice of oligonucleotides for priming the PCR reaction is critically important for selecting the gene of interest from a DNA library (i.e. cloning the gene). Therefore, it will be considered in some detail below.

44. Such oligo nucleotides ("nt") were chosen to be no shorter than 16 nt and preferably about 20-24 nt for optimal annealing for the PCR method. For

example, the probes used to PCR the gene encoding the protein CAAF1 (Seq ID no. 19), were two sets of multiple DNA sequences (Seq ID no. 8, 9) 20 nt in length, which were derived from the N- and C-terminus protein sequence of the S100 protein CAAF1 (Seq ID no. 19).

45. Three types of oligonucleotide probes were commonly used including: (1) single oligonucleotides of defined sequence; (2) a set of oligonucleotides that had sequences that were highly degenerate; and (3) longer oligonucleotides of lesser degeneracy while said longer oligonucleotides could be used, but they were less useful with PCR. In any of these three cases, hybridization depended highly on the materials used (i.e. length and G/C content of probes and template), the temperature that such probes melt (i.e. dissociate from the template), and the salt concentration for PCR to work successfully. Such conditions and probe design required highly skilled researchers and experimentation through trial and error, so the proper conditions and probes could be obtained for PCR.

46. Category (2) oligonucleotides were used most often when the amino acid sequence of the protein of interest was known. This was the method of choice for the cloning the gene for CAAF1 (Seq. ID no. 19). In this method, a set of nucleotides were prepared, which represented all of the possible sequences that could code for the amino acid sequence of the protein of interest. For CAAF1, as an example, the probes listed in the '832 patent (Seq. ID no. 8 and Seq ID no. 9) did not represent all the possible sequences, but rather, were strategically selected to represent the largest subset, and simultaneously eliminating the largest number of possibilities that would provide erroneous sequences from arising.

47. For a simple example, if the amino acid sequence was 6 residues long, and if 5 codons had 2-fold ambiguity and 1 codon had 3-fold ambiguity, then 2 x 2 x 2 x 2 x 2 x 3 = 96 oligonucleotides (18 nt each) would have to be made, so that there is certainty that the DNA encoding this stretch of nucleotides would be hybridized; a similar set of primers would be designed/synthesized for the C-terminal portion of the protein. In the case of CAAF1, two of the isolated peptides contain leucine residues, and this particular amino acid residue has 6-fold ambiguity with respect the possible codons encoding for leucine. Because of the high degree of ambiguity from these leucine residues, the large number of possible oligonucleotides coding for this peptide becomes prohibitively large. In cases such as these (i.e. with leucine residues involved for example), a strategy, whereby two pools are prepared was generally used. In this strategy, one tries to avoid incorrect codons (i.e. TTT and TTC coding for Phe), rather than getting all the possibilities for coding the leucine residue. Other strategies were also available for other amino acid residues that have multiple coding sequences (i.e. those with 4-fold ambiguity such as nucleotides coding for valine, alanine, proline, glycine, threonine). In addition to getting the largest number of probes possible, it was also necessary to optimize reaction conditions (i.e. temperature, salt concentrations etc) especially when all of the possible oligonucleotides could not be synthesized.

48. However, even when there was a full compliment of oligonucleotides, there was still a high chance that isolation of other clones containing similar sequences, but not identical sequences, would be amplified. Specifically, because S100 proteins are highly homologous in one region and have little or no homology in other regions, it was not surprising if one cloned a new S100 protein using primers designed from another S100 protein. Furthermore, one could not predict a

priori from the experimental design whether or not a new S100 protein would be cloned in cases such as these.

49. If a sequence was not known, as was the case here, then it was not possible to isolate this single unknown protein in a predictable way from a cDNA library. Pools of degenerate oligonucleotides were typically used, not to identify target clones unambiguously, but to select a series of clones, which would then require further testing.

50. Multiple tracks of known amino acid sequences were also used to design oligonucleotides with the goal to get a clone of desired sequence; however, without such an amino acid sequence for the entire protein, which was the case with Seq. ID. No. 19, it was not possible to predict *a priori* the sequence to be derived.

51. Screening methods to identify the gene of interest were numerous in 1995, but the final verification was sequencing the DNA itself, once it was inserted into a vector. It was absolutely necessary to sequence the gene inserted into the vector as final verification. However, it should be stated that inserting a gene of interest (i.e. DNA encoding Seq. ID No. 20) into a vector would not enable an ordinary skilled worker in 1995 to have a reasonable chance of success in arriving at Seq. ID No. 19 with the knowledge of Seq. ID No. 20.

52. Once the gene was verified, the methodology to insert its DNA into vectors including prokaryotic and eukaryotic vectors was done in a coordinated manner in concert with deciding DNA sequence that regulates expression and the host cells needed for such expression of the protein. Insertion of a gene into a recombinant expression plasmid, with regulating elements, was usually considered sub-cloning and there were a very large number of strategies in place to do this step in 1995. These

steps required an expert in the field to design such a protein expression system and numerous experiments to work out the details at each step of the sub-cloning and cellular transformation (i.e. prokaryotic and eukaryotic cells including bacteria, yeast, filamentous fungus, plant, and/or animal cells). Methodologies to express and purify proteins expressed from host cells were numerous in 1995, so they will not be reviewed fully here. Usually, a series of experiments were planned, and trial and error methods were used to determine the optimal method to express and purify a recombinant protein from the host cells. However, it should be stated that purifying a recombinant protein of interest (i.e. DNA encoding Seq. ID No. 20) in cells would not enable an ordinary skilled worker in 1995 to have a reasonable chance of success in arriving at Seq. ID No. 19 with the knowledge of Seq. ID No. 20 or vice versa.

53. In 1995, it was not possible to predict exactly what S100 protein sequence(s) you would obtain in a cloning experiment if you only had the amino acid sequence of another S100 protein in hand. Because this S100 protein family has homologous regions it was likely that one would obtain multiple S100 proteins sequences in such 'cloning' experiments as they were performed in 1995, and the exact sequences for the DNA discovered and eventually the amino acid sequences discovered in such experiments were not predictable.

54. The question presented here is relatively straightforward since in 1995 the entire human genome was not yet sequenced, and at that time, the general method to clone a gene was to (1) identify, isolate, and purify a gene product (i.e. protein); (2) sequence the purified protein using standard protein sequencing techniques; (3) create a cDNA library from a tissue from which the protein of interest

was likely expressed – i.e. start with the tissue that one originally isolated the protein from in the first place; (4) extract RNA from the tissue as was necessary to obtain a template from which to synthesize cDNA via reverse transcription procedures; (5) prepare multiple probes to the N- and C-terminal regions of the DNA sequence – a single probe was not sufficient since multiple nucleotide sequences can code for a single amino acid; therefore, such probes are a set mixture that includes each of the 3 nucleotide repeats for a given amino acid residue. Therefore, in such a mixture, at least one probe will be able to anneal to the N- and C-terminus of the gene, respectively, and the DNA encoding this gene can be amplified using polymerase chain reaction (PCR) methodology.

55. While it was possible to amplify other genes in a homologous family when knowledge of the sequence homology is known (i.e. if such probes match or nearly match another gene), and if the correct conditions were selected (i.e. annealing temperature, buffer conditions etc), this would only occur if the two genes were similar in sequence in the N- and C-terminus; whereby, the probes were designed to anneal. Thus, such probes could be used, in principle, to "fish out" genes of similar sequence (i.e. in the area(s) overlapping the probe).

56. It was not possible in 1995 to derive only one gene of interest with a specific DNA sequence without limiting the number of possibilities, which could be done if the protein sequence was known.

57. For example, it would be possible, in the case of S100 proteins, to discover new S100 proteins by having probes that may anneal to mRNA derived from different S100 proteins; however, it would not be possible *a priori* to determine the exact sequence ahead of time. This is an inherent limitation since the methodology required

prior knowledge of the exact sequence to be fully sure that one obtains a specific gene with a specific DNA and protein sequence.

58. For example, if one did not know the sequence of one S100 protein (e.g. seq ID no. 19), but did know the sequence of another S100 protein (e.g. seq ID no. 20), then oligonucleotide probes designed to the N- and C-terminus of these two proteins would be similar and could potentially amplify either gene and perhaps other genes that were not yet known.

59. In other words, because there are homologous DNA regions in homologous proteins, the number of independent perfect matches expected for an oligonucleotide of length "L" in a genome of complexity "C" does not follow the simple equation $P = (1/4)^L \times 2C$, which assumes randomness. This is because the distribution of nucleotides in the coding sequences of mammalian genomes is non-random, partly due to the existence of homologous proteins. Thus, it is likely that for a family of proteins with regions of sequence identity and other regions lacking sequence homology that oligonucleotides used as primers would amplify multiple genes and not just a single gene.

60. Taken that a set of oligonucleotide primers can anneal to multiple genes even under perfect and stringent annealing conditions, the second question is a relatively simple one to answer with straightforward probabilities. That is, what was the chance that an average skilled worker in 1995 could predict the sequence of the variable regions of a protein when the gene was PCR amplified by designing oligonucleotide primers to its homologous region? When seq. ID no. 19 and seq ID no. 20 are compared, of the 91 residues, there are 60 residues that match and 31 residues that are different; most of the residues that are variable are located in two regions of the

protein important for its biological specificity. Therefore, the chance that one could predict how PCR oligonucleotides designed using e.g. Seq ID no. 20 would generate only Seq. ID no. 19 or vice versa is prohibitively low (i.e. 1 chance in 20^{31}) since there are 31 amino acid differences when Seq ID no. 19 and Seq ID no. 2 are compared and there are 20 possible amino acids that could occupy those 31 positions. The chances of predicting the DNA are far less than that for predictions of a protein because there are 3 nucleotides per amino acid residue and multiple codons can code for a single amino acid.

I therefore conclude that the ordinary skilled worker in 1995 would not have a reasonable chance of success in arriving at Seq. ID. No. 19 based upon the knowledge of Seq. No. ID 20 and known procedures or vice versa.

For the reasons described above I have concluded that Seq. ID no. 20 and Seq. ID no. 19 are distinct proteins and likewise their nucleic acid sequence are distinct and that they are novel and non-obvious over one another.

61. In signing this declaration I understand that the declaration will be filed as evidence in a contested case before the Board of Patent Appeals and Interferences of the United States Patent and Trademark Office. Moreover, I acknowledge that I may be subject to cross examination in the case and the cross examination will take place within the United States. If any testimony (deposition, direct testimony, or cross-examination) is required of me, then I will appear provided that I have previously been paid in full for my services and expenses by Anderson Kill & Olick, P.C. and further provided that Anderson Kill & Olick, P.C., pays me in advance, at our agreed upon hourly rates, as well as for travel and incidental expenses, necessary for me to appear. I will appear for cross examination within the United States during the

time allotted for cross examination if consulted ahead of time with regard to my schedule and if I am available.

I hereby declare under penalty of perjury that the foregoing is true and correct.

David J. Weber

Date: January 31, 2007

David J. Weber